Numerical analysis of nonlinear eigenvalue problems

Eric Cancès, Rachida Chakir, Yvon Maday

Numerical analysis of nonlinear eigenvalue problems

Eric Cancès[*], Rachida Chakir[†] and Yvon Maday[†‡]

May 13, 2009

Abstract

We provide *a priori* error estimates for variational approximations of the ground state eigenvalue and eigenvector of nonlinear elliptic eigenvalue problems of the form $-\mathrm{div}(A\nabla u) + Vu + f(u^2)u = \lambda u$, $\|u\|_{L^2} = 1$. We focus in particular on the Fourier spectral approximation (for periodic problems) and on the $\mathbb{P}_1$ and $\mathbb{P}_2$ finite-element discretizations. Denoting by $(u_\delta, \lambda_\delta)$ a variational approximation of the ground state eigenpair (u, λ), we are interested in the convergence rates of $\|u_\delta - u\|_{H^1}$, $\|u_\delta - u\|_{L^2}$ and $|\lambda_\delta - \lambda|$, when the discretization parameter δ goes to zero. We prove that if A, V and f satisfy certain conditions, $|\lambda_\delta - \lambda|$ goes to zero as $\|u_\delta - u\|_{H^1}^2 + \|u_\delta - u\|_{L^2}$. We also show that under more restrictive assumptions on A, V and f, $|\lambda_\delta - \lambda|$ converges to zero as $\|u_\delta - u\|_{H^1}^2$, thus recovering a standard result for *linear* elliptic eigenvalue problems. For the latter analysis, we make use of estimates of the error $u_\delta - u$ in negative Sobolev norms.

1 Introduction

Many mathematical models in science and engineering give rise to nonlinear eigenvalue problems. Let us mention for instance the calculation of the vibration modes of a mechanical structure in the framework of nonlinear elasticity, the Gross-Pitaevskii equation describing the steady states of Bose-Einstein condensates [9], or the Hartree-Fock and Kohn-Sham equations used to calculate ground state electronic structures of molecular systems in quantum chemistry and materials science (see [3] for a mathematical introduction).

The numerical analysis of *linear* eigenvalue problems has been thoroughly studied in the past decades (see e.g. [1]). On the other hand, *nonlinear* eigenvalue problems seem to have received much less attention from numerical analysts.

In this article, we focus on a particular class of nonlinear eigenvalue problems arising in the study of variational models of the form

$$I = \inf\left\{E(v),\ v \in X,\ \int_\Omega v^2 = 1\right\} \tag{1}$$

where

Ω is a regular bounded domain or a rectangular brick of $\mathbb{R}^d$ and $X = H_0^1(\Omega)$
or
Ω is the unit cell of a periodic lattice $\mathcal{R}$ of $\mathbb{R}^d$ and $X = H_\#^1(\Omega)$

[*]Université Paris-Est, CERMICS, Project-team Micmac, INRIA-Ecole des Ponts, 6 & 8 avenue Blaise Pascal, 77455 Marne-la-Vallée Cedex 2, France.
[†]UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France
[‡]Division of Applied Mathematics, Brown University, Providence, RI, USA

with $d = 1$, 2 or 3, and where the energy functional E is of the form

$$E(v) = \frac{1}{2}a(v,v) + \frac{1}{2}\int_\Omega F(v^2(x))\,dx \tag{2}$$

with

$$a(u,v) = \int_\Omega (A\nabla u)\cdot\nabla v + \int_\Omega Vuv. \tag{3}$$

Recall that if Ω is the unit cell of a periodic lattice $\mathcal{R}$ of $\mathbb{R}^d$, then for all $s \in \mathbb{R}$ and $k \in \mathbb{N}$,

$$\begin{aligned} H^s_\#(\Omega) &= \left\{v|_\Omega,\ v \in H^s_{\rm loc}(\mathbb{R}^d) \mid v\ \mathcal{R}\text{-periodic}\right\}, \\ C^k_\#(\Omega) &= \left\{v|_\Omega,\ v \in C^k(\mathbb{R}^d) \mid v\ \mathcal{R}\text{-periodic}\right\}. \end{aligned}$$

We assume in addition that

- $A \in (L^\infty(\Omega))^{d\times d}$ and $A(x)$ is symmetric for almost all $x \in \Omega$ (4)
  $\exists \alpha > 0$ s.t. $\xi^T A(x)\xi \geq \alpha|\xi|^2$ for all $\xi \in \mathbb{R}^d$ and almost all $x \in \Omega$ (5)

- $V \in L^p(\Omega)$ for some $p > 2$ (6)

- $F \in C^1([0,+\infty),\mathbb{R}) \cap C^2((0,\infty),\mathbb{R})$ and $F'' > 0$ on $(0,+\infty)$ (7)
  $\forall R > 0,\ \exists C_R \in \mathbb{R}_+$ s.t. $\forall 0 < t_1 \leq R,\ \forall t_2 \in \mathbb{R}$,
  $\left|(F'(t_2^2) - F'(t_1^2))t_2^2\right| \leq C_R(1+|t_2|^3)|t_2 - t_1|$ and (8)
  $\left|F'(t_2^2)t_2 - F'(t_1^2)t_2 - 2F''(t_1^2)t_1^2(t_2-t_1)\right| \leq C_R(1+|t_2|)|t_2-t_1|^2$. (9)

In particular, the function $F(t) = ct^q$ satisfies the assumptions (7)-(9) if and only if $c > 0$ and $\frac{3}{2} \leq q \leq 2$. This allows us to handle the Thomas-Fermi kinetic energy functional ($q = \frac{5}{3}$) as well as the repulsive interaction in Bose-Einstein condensates ($q = 2$). In order to simplify the notation, we denote by $f(t) = F'(t)$.

Making the change of variable $\rho = v^2$ and noticing that $a(|v|,|v|) = a(v,v)$ for all $v \in X$, it is easy to check that

$$I = \inf\left\{\mathcal{E}(\rho),\ \rho \geq 0,\ \sqrt{\rho} \in X,\ \int_\Omega \rho = 1\right\}, \tag{10}$$

where

$$\mathcal{E}(\rho) = \frac{1}{2}a(\sqrt{\rho},\sqrt{\rho}) + \frac{1}{2}\int_\Omega F(\rho).$$

Under assumptions (4)-(9), (10) has a unique solution ρ_0 and (1) has exactly two solutions: $u = \sqrt{\rho_0}$ and $-u$. Besides, E is C^1 on X and for all $v \in X$, $E'(v) = A_v v$ where

$$A_v = -\mathrm{div}(A\nabla\cdot) + V + f(v^2).$$

Note that A_v defines a self-adjoint operator on $L^2(\Omega)$, with form domain X. The function u therefore is solution to the Euler equation

$$\forall v \in X, \quad \langle E'(u) - \lambda u, v\rangle_{X',X} = 0 \tag{11}$$

for some $\lambda \in \mathbb{R}$ (the Lagrange multiplier of the constraint $\|u\|^2_{L^2} = 1$) and equation (11), complemented with the constraint $\|u\|_{L^2} = 1$, takes the form of the nonlinear eigenvalue problem

$$\left\{\begin{array}{l} A_u u = \lambda u \\ \|u\|_{L^2} = 1. \end{array}\right. \tag{12}$$

In addition, $u \in C^0(\overline{\Omega})$, $u > 0$ in Ω and λ is the ground state eigenvalue of the linear operator A_u. An important result is that λ is a *simple* eigenvalue of A_u. All these properties are classical. For the sake of completeness, their proofs are however given in the Appendix.

Let us now turn to the main topic of this article, namely the derivation of a priori error estimates for variational approximations of the ground state eigenpair (λ, u). We denote by $(X_\delta)_{\delta>0}$ a family of finite-dimensional subspaces of X such that

$$\min\{\|u - v_\delta\|_{H^1},\ v_\delta \in X_\delta\} \underset{\delta\to 0^+}{\longrightarrow} 0 \tag{13}$$

and consider the variational approximation of (1) consisting in solving

$$I_\delta = \inf\left\{E(v_\delta),\ v_\delta \in X_\delta,\ \int_\Omega v_\delta^2 = 1\right\}. \tag{14}$$

Problem (14) has at least one minimizer u_δ, which satisfies

$$\forall v_\delta \in X_\delta, \quad \langle E'(u_\delta) - \lambda_\delta u_\delta, v_\delta\rangle_{X',X} = 0 \tag{15}$$

for some $\lambda_\delta \in \mathbb{R}$. Obviously, $-u_\delta$ also is a minimizer associated with the same eigenvalue λ_δ. On the other hand, it is not known whether u_δ and $-u_\delta$ are the only minimizers of (14). This follows from the fact that the set

$$\left\{\rho \mid \exists u_\delta \in X_\delta \text{ s.t. } \|u_\delta\|_{L^2} = 1,\ \rho = u_\delta^2\right\}$$

is not convex in general. We will see however (cf. Theorem 1) that for any global minimum u_δ of (14) such that $(u, u_\delta) \geq 0$, the following holds true

$$\|u_\delta - u\|_{H^1} \underset{\delta\to 0^+}{\longrightarrow} 0.$$

In addition, a simple calculation leads to

$$\lambda_\delta - \lambda = \langle (A_u - \lambda)(u_\delta - u), (u_\delta - u)\rangle_{X',X} + \int_\Omega w_{u,u_\delta}(u_\delta - u) \tag{16}$$

where

$$w_{u,u_\delta} = u_\delta^2 \frac{f(u_\delta^2) - f(u^2)}{u_\delta - u}.$$

The first term of the right-hand side of (16) is nonnegative and goes to zero as $\|u_\delta - u\|_{H^1}^2$. We will prove in Theorem 1 that the second term goes to zero at least as $\|u_\delta - u\|_{L^2}$. Therefore, $|\lambda_\delta - \lambda|$ converges to zero with δ at least as $\|u_\delta - u\|_{H^1}^2 + \|u_\delta - u\|_{L^2}$.

The purpose of this article is to provide more precise *a priori* error bounds on $\|u_\delta - u\|_{H^1}$, $\|u_\delta - u\|_{L^2}$ and $|\lambda_\delta - \lambda|$. In Section 2, we prove a series of estimates valid in the general framework described above. We then turn to more specific examples, where the analysis can be pushed further. In Section 3, we concentrate on the discretization of problem (1) with

$$\begin{aligned}
&\Omega = (0, 2\pi)^d,\\
&X = H^1_\#(0, 2\pi)^d,\\
&E(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 + \frac{1}{2}\int_\Omega V v^2 + \frac{1}{2}\int_\Omega F(v^2),
\end{aligned}$$

in Fourier modes. In Section 4, we deal with the $\mathbb{P}_1$ and $\mathbb{P}_2$ finite element discretizations of problem (1) with

$$\begin{aligned}&\Omega \text{ rectangular brick of } \mathbb{R}^d,\\ &X = H_0^1(\Omega),\\ &E(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 + \frac{1}{2}\int_\Omega V v^2 + \frac{1}{2}\int_\Omega F(v^2).\end{aligned}$$

Lastly, we discuss the issue of numerical integration in Section 5.

2 Basic error analysis

The aim of this section is to establish error bounds on $\|u_\delta - u\|_{H^1}$, $\|u_\delta - u\|_{L^2}$ and $|\lambda_\delta - \lambda|$ in a general framework. In the whole section, we make the assumptions (4)-(9) and (13), and we denote by u the unique positive solution of (1) and by u_δ a minimizer of the discretized problem (14) such that $(u_\delta, u)_{L^2} \geq 0$.

Lemma 1 *The functional E is twice differentiable at u and for all $(v, w) \in X \times X$,*

$$\langle E''(u)v, w\rangle = \langle A_u v, w\rangle_{X',X} + 2\int_\Omega f'(u^2)u^2 vw. \tag{17}$$

There exists $\beta > 0$ and $M \in \mathbb{R}_+$ such that for all $v \in X$,

$$0 \leq \langle (A_u - \lambda)v, v\rangle_{X',X} \leq M\|v\|_{H^1}^2 \tag{18}$$

$$\beta\|v\|_{H^1}^2 \leq \langle (E''(u) - \lambda)v, v\rangle_{X',X} \leq M\|v\|_{H^1}^2. \tag{19}$$

There exists $\gamma > 0$ such that for all $\delta > 0$,

$$\gamma\|u_\delta - u\|_{H^1}^2 \leq \langle (A_u - \lambda)(u_\delta - u), (u_\delta - u)\rangle_{X',X}. \tag{20}$$

Proof The quadratic functional $v \mapsto \frac{1}{2}a(v, v)$ is clearly twice differentiable on X. Using (8) and (9), it is easy to check that the functional $\Phi : v \mapsto \frac{1}{2}\int_\Omega F(v^2)$ is twice differentiable on X as well and that

$$\Phi'(u) = f(u^2)u, \qquad \langle \Phi''(u)v, w\rangle_{X',X} = \int_\Omega (f(u^2) + 2f'(u^2)u^2)vw.$$

This straightforwardly leads to (17). We have for all $v \in X$,

$$\langle (A_u - \lambda)v, v\rangle_{X',X} \leq \|A\|_{L^\infty}\|\nabla v\|_{L^2}^2 + \|V\|_{L^2}\|v\|_{L^4}^2 + \|f(u^2)\|_{L^\infty}\|v\|_{L^2}^2$$

and

$$\langle (E''(u) - \lambda)v, v\rangle_{X',X} \leq \langle (A_u - \lambda)v, v\rangle_{X',X} + 2\|f'(u^2)u^2\|_{L^\infty}\|v\|_{L^2}^2.$$

Hence the upper bounds in (18) and (19). We now use the fact that λ, the lowest eigenvalue of A_u, is simple (see Lemma 2 in the Appendix). This implies that there exists $\eta > 0$ such that

$$\forall v \in X, \quad \langle (A_u - \lambda)v, v\rangle_{X',X} \geq \eta(\|v\|_{L^2}^2 - |(u, v)_{L^2}|^2) \geq 0. \tag{21}$$

This provides on the one hand the lower bound (18), and leads on the other hand to the inequality

$$\forall v \in X, \quad \langle (E''(u) - \lambda)v, v\rangle_{X',X} \geq 2\int_\Omega f'(u^2)u^2v^2.$$

As $f' = F'' > 0$ in $(0, +\infty)$ and $u > 0$ in Ω, we therefore have

$$\forall v \in X \setminus \{0\}, \quad \langle (E''(u) - \lambda)v, v \rangle_{X',X} > 0.$$

Reasoning by contradiction, we deduce from the above inequality and the first inequality in (21) that there exists $\widetilde{\eta} > 0$ such that

$$\forall v \in X, \quad \langle (E''(u) - \lambda)v, v \rangle_{X',X} \geq \widetilde{\eta} \|v\|_{L^2}^2. \tag{22}$$

Besides, there exists a constant $C \in \mathbb{R}_+$ such that

$$\forall v \in X, \quad \langle (A_u - \lambda)v, v \rangle_{X',X} \geq \frac{\alpha}{2} \|\nabla v\|_{L^2}^2 - C\|u\|_{L^2}^2. \tag{23}$$

Let us establish this inequality for $d = 3$ (the case when $d = 1$ is straightforward and the case when $d = 2$ can be dealt with in the same way). For all $x \in X$,

$$\begin{aligned}
\langle (A_u - \lambda)v, v \rangle_{X',X} &= \int_\Omega (A\nabla v) \cdot \nabla v + \int_\Omega (V + f(v^2) - \lambda)v^2 \\
&\geq \alpha \|\nabla v\|_{L^2}^2 - \|V\|_{L^2} \|v\|_{L^4}^2 + (f(0) - \lambda)\|v\|_{L^2}^2 \\
&\geq \alpha \|\nabla v\|_{L^2}^2 - \|V\|_{L^2} \|v\|_{L^2}^{1/2} \|v\|_{L^6}^{3/2} + (f(0) - \lambda)\|v\|_{L^2}^2 \\
&\geq \alpha \|\nabla v\|_{L^2}^2 - C_6^{3/2} \|V\|_{L^2} \|v\|_{L^2}^{1/2} \|\nabla v\|_{L^6}^{3/2} + (f(0) - \lambda)\|v\|_{L^2}^2 \\
&\geq \frac{\alpha}{2} \|\nabla v\|_{L^2}^2 + \left(f(0) - \lambda - \frac{27 C_6^6}{32 \alpha^3} \|V\|_{L^2}^4 \right) \|v\|_{L^2}^2,
\end{aligned}$$

where C_6 is the Sobolev constant such that $\forall v \in X$, $\|v\|_{L^6} \leq C_6 \|\nabla v\|_{L^2}$. The coercivity of $E''(u) - \lambda$ (i.e. the lower bound in (19)) is a straightforward consequence of (22) and (23).

To prove (20), we notice that

$$\|u_\delta\|_{L^2}^2 - |(u, u_\delta)_{L^2}|^2 \geq 1 - (u, u_\delta)_{L^2} = \frac{1}{2} \|u_\delta - u\|_{L^2}^2.$$

It therefore readily follows from (21) that

$$\langle (A_u - \lambda)(u_\delta - u), (u_\delta - u) \rangle_{X',X} \geq \frac{\eta}{2} \|u_\delta - u\|_{L^2}^2.$$

Combining with (23), we finally obtain (20). □

For $w \in X'$, we denote by ψ_w the unique solution to the adjoint problem

$$\left\{ \begin{array}{l} \text{find } \psi_w \in u^\perp \text{ such that} \\ \forall v \in u^\perp, \quad \langle (E''(u) - \lambda)\psi_w, v \rangle_{X',X} = \langle w, v \rangle_{X',X}, \end{array} \right. \tag{24}$$

where

$$u^\perp = \left\{ v \in X \mid \int_\Omega uv = 0 \right\}.$$

The existence and uniqueness of the solution to (24) is a straightforward consequence of (19) and the Lax-Milgram lemma. Besides,

$$\forall w \in X', \quad \|\psi_w\|_{H^1} \leq \beta^{-1} M \|w\|_{X'} \leq \beta^{-1} M \|w\|_{L^2}. \tag{25}$$

We can now state the main result of this section.

Theorem 1 *It holds*

$$\|u_\delta - u\|_{H^1} \underset{\delta\to 0^+}{\longrightarrow} 0. \tag{26}$$

Besides, there exists $\delta_0 > 0$ *and* $C \in \mathbb{R}_+$ *such that for all* $0 < \delta < \delta_0$,

$$\begin{aligned}
\|u_\delta - u\|_{H^1} &\leq C \min_{v_\delta \in X_\delta} \|v_\delta - u\|_{H^1} && (27)\\
\|u_\delta - u\|_{L^2}^2 &\leq C\Big(\|u_\delta - u\|_{H^1}^2 \|u_\delta - u\|_{L^2} \\
&\qquad + \|u_\delta - u\|_{H^1} \min_{\psi_\delta \in X_\delta} \|\psi_{u_\delta - u} - \psi_\delta\|_{H^1} \Big) && (28)\\
|\lambda_\delta - \lambda| &\leq C\left(\|u_\delta - u\|_{H^1}^2 + \|u_\delta - u\|_{L^2} \right). && (29)
\end{aligned}$$

Proof Using (20) and the convexity of F, we get

$$\begin{aligned}
E(u_\delta) - E(u) &= \frac{1}{2}\langle A_u u_\delta, u_\delta\rangle_{X',X} - \frac{1}{2}\langle A_u u, u\rangle_{X',X} \\
&\quad + \frac{1}{2}\int_\Omega F\left(u_\delta^2\right) - F\left(u^2\right) - f\left(u^2\right)\left(u_\delta^2 - u^2\right) \\
&= \frac{1}{2}\langle (A_u - \lambda)(u_\delta - u), (u_\delta - u)\rangle_{X',X} \\
&\quad + \frac{1}{2}\int_\Omega F\left(u^2 + (u_\delta^2 - u^2)\right) - F\left(u^2\right) - f\left(u^2\right)\left(u_\delta^2 - u^2\right) \\
&\geq \frac{\gamma}{2}\|u_\delta - u\|_{H^1}^2. && (30)
\end{aligned}$$

Let $\Pi_\delta u \in X_\delta$ be such that

$$\|u - \Pi_\delta u\|_{H^1} = \min\left\{\|u - v_\delta\|_{H^1}, v_\delta \in X_\delta\right\}.$$

We deduce from (13) that $(\Pi_\delta u)_{\delta>0}$ converges to u in X when δ goes to zero. Denoting by $\widetilde{u}_\delta = \|\Pi_\delta u\|_{L^2}^{-1}\Pi_\delta u$ (which is well defined, at least for δ small enough), we also have

$$\lim_{\delta\to 0^+} \|\widetilde{u}_\delta - u\|_{H^1} = 0.$$

The functional E being strongly continuous on X, we obtain

$$\|u_\delta - u\|_{H^1}^2 \leq \frac{2}{\gamma}\left(E(u_\delta) - E(u)\right) \leq \frac{2}{\gamma}\left(E(\widetilde{u}_\delta) - E(u)\right) \underset{\delta\to 0^+}{\longrightarrow} 0. \tag{31}$$

It follows that there exists $\delta_1 > 0$ such that

$$\forall 0 < \delta \leq \delta_1, \quad \|u_\delta\|_{H^1} \leq 2\|u\|_{H^1}, \quad \|u_\delta - u\|_{H^1} \leq \frac{1}{2}.$$

Next, we remark that

$$\begin{aligned}
\lambda_\delta - \lambda &= \langle E'(u_\delta), u_\delta\rangle_{X',X} - \langle E'(u), u\rangle_{X',X} \\
&= a(u_\delta, u_\delta) - a(u,u) + \int_\Omega f(u_\delta^2)u_\delta^2 - \int_\Omega f(u^2)u^2 \\
&= a(u_\delta - u, u_\delta - u) + 2a(u, u_\delta - u) + \int_\Omega f(u_\delta^2)u_\delta^2 - \int_\Omega f(u^2)u^2 \\
&= a(u_\delta - u, u_\delta - u) + 2\lambda \int_\Omega u(u_\delta - u) - 2\int_\Omega f(u^2)u(u_\delta - u) \\
&\qquad + \int_\Omega f(u_\delta^2)u_\delta^2 - \int_\Omega f(u^2)u^2 \\
&= a(u_\delta - u, u_\delta - u) - \lambda\|u_\delta - u\|_{L^2}^2 - 2\int_\Omega f(u^2)u(u_\delta - u) \\
&\qquad + \int_\Omega f(u_\delta^2)u_\delta^2 - \int_\Omega f(u^2)u^2 \\
&= \langle (A_u - \lambda)(u_\delta - u), (u_\delta - u)\rangle_{X',X} + \int_\Omega w_{u,u_\delta}(u_\delta - u) \qquad (32)
\end{aligned}$$

where

$$w_{u,u_\delta} = u_\delta^2 \frac{f(u_\delta^2) - f(u^2)}{u_\delta - u}.$$

Using (8) and (18), we obtain that for all $0 < \delta \le \delta_1$,

$$\begin{aligned}
|\lambda_\delta - \lambda| &\le M\|u_\delta - u\|_{H^1}^2 + \|w_{u,u_\delta}\|_{L^2}\|u - u_\delta\|_{L^2} \\
&\le M\|u_\delta - u\|_{H^1}^2 + C\|1 + |u_\delta|^3\|_{L^2}\|u - u_\delta\|_{L^2} \\
&\le M\|u_\delta - u\|_{H^1}^2 + C(1 + \|u_\delta\|_{H^1}^3)\|u - u_\delta\|_{L^2} \\
&\le C\left(\|u_\delta - u\|_{H^1}^2 + \|u - u_\delta\|_{L^2}\right), \qquad (33)
\end{aligned}$$

where C denotes constants independent of δ.

In order to evaluate the H^1-norm of the error $u_\delta - u$, we first notice that

$$\forall v_\delta \in X_\delta, \quad \|u_\delta - u\|_{H^1} \le \|u_\delta - v_\delta\|_{H^1} + \|v_\delta - u\|_{H^1}, \qquad (34)$$

and that

$$\begin{aligned}
\|u_\delta - v_\delta\|_{H^1}^2 &\le \beta^{-1}\langle (E''(u) - \lambda)(u_\delta - v_\delta), (u_\delta - v_\delta)\rangle_{X',X} \\
&= \beta^{-1}\Big(\langle (E''(u) - \lambda)(u_\delta - u), (u_\delta - v_\delta)\rangle_{X',X} \\
&\qquad + \langle (E''(u) - \lambda)(u - v_\delta), (u_\delta - v_\delta)\rangle_{X',X}\Big). \qquad (35)
\end{aligned}$$

For all $w_\delta \in X_\delta$

$$\begin{aligned}
&\langle (E''(u) - \lambda)(u_\delta - u), w_\delta\rangle_{X',X} \\
&\quad = -\int_\Omega \left(f(u_\delta^2)u_\delta - f(u^2)u_\delta - 2f'(u^2)u^2(u_\delta - u)\right) w_\delta + (\lambda_\delta - \lambda)\int_\Omega (u_\delta - u)w_\delta.
\end{aligned}$$

Using (9) and (33), we therefore obtain that for all $0 < \delta \le \delta_1$ and all $w_\delta \in X_\delta$,

$$|\langle (E''(u) - \lambda)(u_\delta - u), w_\delta\rangle_{X',X}| \le C\|w_\delta\|_{H^1}\|u_\delta - u\|_{H^1}^2. \qquad (36)$$

It then follows from (19), (35) and (36) that for all $0 < \delta \le \delta_1$ and all $v_\delta \in X_\delta$,

$$\|u_\delta - v_\delta\|_{H^1} \le C\left(\|v_\delta - u\|_{H^1} + \|u_\delta - u\|_{H^1}^2\right).$$

Combining with (34) we obtain that there exists $0 < \delta_2 \leq \delta_1$ and $C \in \mathbb{R}_+$ such that for all $0 < \delta \leq \delta_2$ and all $v_\delta \in X_\delta$,

$$\|u_\delta - u\|_{H^1} \leq C\|v_\delta - u\|_{H^1}. \tag{37}$$

Thus (27) is proved.

Let u_δ^* be the orthogonal projection, for the L^2 inner product, of u_δ on the affine space $\left\{v \in L^2(\Omega) \,|\, \int_\Omega uv = 1\right\}$. One has

$$u_\delta^* \in X, \qquad u_\delta^* - u \in u^\perp, \qquad u_\delta^* - u_\delta = \frac{1}{2}\|u_\delta - u\|_{L^2}^2 u,$$

from which we infer that

$$\begin{aligned}
\|u_\delta - u\|_{L^2}^2 &= \int_\Omega (u_\delta - u)(u_\delta^* - u) + \int_\Omega (u_\delta - u)(u_\delta - u_\delta^*) \\
&= \int_\Omega (u_\delta - u)(u_\delta^* - u) - \frac{1}{2}\|u_\delta - u\|_{L^2}^2 \int_\Omega (u_\delta - u)u \\
&= \int_\Omega (u_\delta - u)(u_\delta^* - u) + \frac{1}{2}\|u_\delta - u\|_{L^2}^2 \left(1 - \int_\Omega u_\delta u\right) \\
&= \int_\Omega (u_\delta - u)(u_\delta^* - u) + \frac{1}{4}\|u_\delta - u\|_{L^2}^4 \\
&= \langle u_\delta - u, u_\delta^* - u\rangle_{X',X} + \frac{1}{4}\|u_\delta - u\|_{L^2}^4 \\
&= \langle (E''(u) - \lambda)\psi_{u_\delta - u}, u_\delta^* - u\rangle_{X',X} + \frac{1}{4}\|u_\delta - u\|_{L^2}^4 \\
&= \langle (E''(u) - \lambda)(u_\delta - u), \psi_{u_\delta - u}\rangle_{X',X} \\
&\quad + \frac{1}{2}\|u_\delta - u\|_{L^2}^2 \langle (E''(u) - \lambda)u, \psi_{u_\delta - u}\rangle_{X',X} + \frac{1}{4}\|u_\delta - u\|_{L^2}^4 \\
&= \langle (E''(u) - \lambda)(u_\delta - u), \psi_{u_\delta - u}\rangle_{X',X} \\
&\quad + \|u_\delta - u\|_{L^2}^2 \int_\Omega f'(u^2)u^3\psi_{u_\delta - u} + \frac{1}{4}\|u_\delta - u\|_{L^2}^4.
\end{aligned}$$

For all $\psi_\delta \in X_\delta$, it holds

$$\begin{aligned}
\|u_\delta - u\|_{L^2}^2 &= \langle (E''(u) - \lambda)(u_\delta - u), \psi_\delta\rangle_{X',X} \\
&\quad + \langle (E''(u) - \lambda)(u_\delta - u), \psi_{u_\delta - u} - \psi_\delta\rangle_{X',X} \\
&\quad + \|u_\delta - u\|_{L^2}^2 \int_\Omega f'(u^2)u^3\psi_{u_\delta - u} + \frac{1}{4}\|u_\delta - u\|_{L^2}^4 \\
&= -\int_\Omega \left(f(u_\delta^2)u_\delta - f(u^2)u_\delta - 2f'(u^2)u^2(u_\delta - u)\right)\psi_\delta \\
&\quad + (\lambda_\delta - \lambda)\int_\Omega u_\delta\psi_\delta \\
&\quad + \langle (E''(u) - \lambda)(u_\delta - u), \psi_{u_\delta - u} - \psi_\delta\rangle_{X',X} \\
&\quad + \|u_\delta - u\|_{L^2}^2 \int_\Omega f'(u^2)u^3\psi_{u_\delta - u} + \frac{1}{4}\|u_\delta - u\|_{L^2}^4 \\
&= -\int_\Omega \left(f(u_\delta^2)u_\delta - f(u^2)u_\delta - 2f'(u^2)u^2(u_\delta - u)\right)\psi_\delta \\
&\quad + (\lambda_\delta - \lambda)\int_\Omega u_\delta(\psi_\delta - \psi_{u_\delta - u}) + (\lambda_\delta - \lambda)\int_\Omega (u_\delta - u)\psi_{u_\delta - u} \\
&\quad + \langle (E''(u) - \lambda)(u_\delta - u), \psi_{u_\delta - u} - \psi_\delta\rangle_{X',X} \\
&\quad + \|u_\delta - u\|_{L^2}^2 \int_\Omega f'(u^2)u^3\psi_{u_\delta - u} + \frac{1}{4}\|u_\delta - u\|_{L^2}^4,
\end{aligned}$$

where we have used that $\displaystyle\int_\Omega \psi_{u_\delta - u} u = 0$.

Let $\psi_\delta^0 \in X_\delta$ be such that

$$\|\psi_{u_\delta - u} - \psi_\delta^0\|_{H^1} = \min_{\psi_\delta \in X_\delta} \|\psi_{u_\delta - u} - \psi_\delta\|_{H^1}.$$

Noticing that $\|\psi_\delta^0\|_{H^1} \leq \|\psi_{u_\delta - u}\|_{H^1} \leq \beta^{-1} M \|u_\delta - u\|_{L^2}$, we obtain from (9), (19) and (33) that there exists $C \in \mathbb{R}_+$ such that for all $0 < \delta \leq \delta_1$,

$$\begin{aligned} \|u_\delta - u\|_{L^2}^2 \quad \leq \quad & C\Big(\|u_\delta - u\|_{H^1}^2 \|u_\delta - u\|_{L^2} + \|u_\delta - u\|_{H^1} \|\psi_{u_\delta - u} - \psi_\delta^0\|_{H^1} \\ & + \|u_\delta - u\|_{L^2}^3 + \|u_\delta - u\|_{L^2}^2 \|u_\delta - u\|_{H^1}^2 \Big). \end{aligned}$$

Therefore, there exists $0 < \delta_0 \leq \delta_2$ and $C \in \mathbb{R}_+$ such that for all $0 < \delta \leq \delta_0$,

$$\|u_\delta - u\|_{L^2}^2 \leq C \left(\|u_\delta - u\|_{H^1}^2 \|u_\delta - u\|_{L^2} + \|u_\delta - u\|_{H^1} \|\psi_{u_\delta - u} - \psi_\delta^0\|_{H^1} \right), \tag{38}$$

which completes the proof of Theorem 1. □

Remark 1 *In the proof of Theorem 1, we have obtained bounds on $|\lambda_\delta - \lambda|$ from (32), using L^2 estimates on w_{u,u_δ} and $(u_\delta - u)$ to control the second term of the right hand side. Remarking that*

$$\begin{aligned} \nabla w_{u,u_\delta} \quad = \quad & -u \, \frac{f(u^2)u - f(u_\delta^2)u - 2f'(u_\delta^2)u_\delta^2(u - u_\delta)}{(u_\delta - u)^2} \, \nabla u_\delta \\ & - u_\delta \, \frac{f(u_\delta^2)u_\delta - f(u^2)u_\delta - 2f'(u^2)u^2(u_\delta - u)}{(u_\delta - u)^2} \, \nabla u \\ & + 2uu_\delta \left(f'(u_\delta^2) \, \nabla u_\delta + f'(u^2) \nabla u \right) + 2u_\delta \, \frac{f(u_\delta^2) - f(u^2)}{u_\delta - u} \, \nabla u_\delta \,, \end{aligned}$$

we deduce from (8)-(9) that if u_δ is uniformly bounded in $L^\infty(\Omega)$, then w_{u,u_δ} is uniformly bounded in X. It then follows from (32) that

$$|\lambda_\delta - \lambda| \leq C \left(\|u_\delta - u\|_{H^1}^2 + \|u_\delta - u\|_{X'} \right),$$

an estimate which is an improvement of (29). In the next two sections, we will see that this approach (or analogous strategies making use of negative Sobolev norms of higher orders), can be used in certain cases to obtain optimal estimates on $|\lambda_\delta - \lambda|$ of the form

$$|\lambda_\delta - \lambda| \leq C \|u_\delta - u\|_{H^1}^2.$$

3 Fourier expansion

In this section, we consider the problem

$$\inf \left\{ E(v), \; v \in X, \; \int_\Omega v^2 = 1 \right\}, \tag{39}$$

where

$$\begin{aligned} & \Omega = (0, 2\pi)^d, \quad \text{with } d = 1, 2 \text{ or } 3, \\ & X = H^1_\#(\Omega), \\ & E(v) = \frac{1}{2} \int_\Omega |\nabla v|^2 + \frac{1}{2} \int_\Omega V v^2 + \frac{1}{2} \int_\Omega F(v^2). \end{aligned}$$

We assume that $V \in H^\sigma_\#(\Omega)$ for some $\sigma > d/2$ and that the function F satisfies (7)-(9) and is such that the function $t \mapsto F(t^2)$ is in $C^{\sigma+1}(\mathbb{R}_+, \mathbb{R})$ if $\sigma \in \mathbb{N}$, and in $C^{[\sigma]+2}(\mathbb{R}_+, \mathbb{R})$ if $\sigma \notin \mathbb{N}$. Actually, as we know that the unique positive solution u to (39) stays away from 0 in Ω, the assumptions on F can be slightly relaxed, but we will not elaborate further on this technical detail.

The positive solution u to (39), which satisfies the elliptic equation

$$-\Delta u + Vu + f(u^2)u = \lambda u,$$

then is in $H^{\sigma+2}_\#(\Omega)$.

A natural discretization of this problem consists in using a Fourier basis. Denoting by $e_k(x) = (2\pi)^{-d/2} e^{ik\cdot x}$, we have for all $v \in L^2(\Omega)$,

$$v(x) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \widehat{v}_k e_k(x),$$

where $\widehat{v}_k$ is the k^{th} Fourier coefficient of v:

$$\widehat{v}_k = \int_\Omega v(x)\, \overline{e_k(x)}\, dx = (2\pi)^{-d/2} \int_\Omega v(x)\, e^{-ik\cdot x}\, dx.$$

The approximation of the solution to (39) by the spectral Fourier approximation is based on the choice

$$X_\delta = \widetilde{X}_N = \text{Span}\{e_k,\ |k|_* \le N\},$$

where $|k|_*$ denotes either the l^2-norm or the l^∞-norm of k (i.e. either $|k| = (\sum_{i=1}^d |k_i|^2)^{1/2}$ or $|k|_\infty = \max_{1\le i\le d} |k_i|$). For convenience, the discretization parameter for this approximation will be denoted as N.

Endowing $H^r_\#(\Omega)$ with the norm defined by

$$\|v\|_{H^r} = \left(\sum_{k \in \mathbb{Z}^d} \left(1 + |k|_*^2\right)^r |\widehat{v}_k|^2 \right)^{1/2},$$

we obtain that for all $s \in \mathbb{R}$, and all $v \in H^s_\#(\Omega)$, the best approximation of v in $H^r_\#(\Omega)$ for any $r \le s$ is

$$\Pi_N v = \sum_{\boldsymbol{k} \in \mathbb{Z}^d, |\boldsymbol{k}|_* \le N} \widehat{v}_k e_k. \tag{40}$$

The more regular v (the regularity being measured in terms of the Sobolev norms H^r), the faster the convergence of this truncated series to v: for all real numbers r and s with $r \le s$, we have

$$\forall v \in H^s_\#(\Omega), \quad \|v - \Pi_N v\|_{H^r} \le \frac{1}{N^{s-r}} \|v\|_{H^s}. \tag{41}$$

Let u_N be a solution to the variational problem

$$\inf \left\{ E(v_N),\ v_N \in \widetilde{X}_N,\ \int_\Omega v_N^2 = 1 \right\} \tag{42}$$

such that $(u_N, u)_{L^2} \ge 0$. Using (41), we obtain

$$\|u - \Pi_N u\|_{H^1} \le \frac{1}{N^{\sigma+1}} \|u\|_{H^{\sigma+2}},$$

Besides, the unique solution to (24) solves the elliptic equation

$$-\Delta\psi_w + \left(V + f(u^2) + 2f'(u^2)u^2 - \lambda\right)\psi_w = 2\left(\int_\Omega f'(u^2)u^3\psi_w\right)u + w - (w,u)_{L^2}u,$$

from which we infer that if $w \in H^r_\#(\Omega)$ for some $0 \le r \le \sigma$, then ψ_w is in $H^{r+2}_\#(\Omega)$ and satisfies

$$\|\psi_w\|_{H^{r+2}} \le C\|w\|_{H^r}, \tag{43}$$

for some constant C independent of w. In particular,

$$\|\psi_{u_N-u} - \Pi_N\psi_{u_N-u}\|_{H^1} \le \frac{1}{N}\|\psi_{u_N-u}\|_{H^2} \le \frac{C}{N}\|u_N - u\|_{L^2}.$$

A direct application of Theorem 1 then yields that there exists $\delta_0 > 0$ and $C \in \mathbb{R}_+$ such that for all $0 < \delta \le \delta_0$,

$$\|u_N - u\|_{H^1} \le \frac{C}{N^{\sigma+1}} \tag{44}$$

$$\|u_N - u\|_{L^2} \le \frac{C}{N^{\sigma+2}} \tag{45}$$

$$|\lambda_N - \lambda| \le \frac{C}{N^{\sigma+2}}. \tag{46}$$

The last estimate is slightly deceptive since, in the case of a linear eigenvalue problem (i.e. for $-\Delta u + Vu = \lambda u$) the convergence of the eigenvalues goes twice as fast as the convergence of the eigenvector in the H^1-norm. We are going to prove that this is also the case for the nonlinear eigenvalue problem under study in this section, at least under some additional assumptions on the function f.

Let us first come back to (32), which we rewrite as,

$$\lambda_N - \lambda = \langle (A_u - \lambda)(u_N - u), (u_N - u)\rangle_{X',X} + \int_\Omega w_{u,u_N}(u_N - u) \tag{47}$$

with

$$w_{u,u_N} = u_N^2\frac{f(u_N^2) - f(u^2)}{u_N - u}.$$

We then observe that u_N is solution to the elliptic equation

$$-\Delta u_N + \Pi_N\left[Vu_N + f(u_N^2)u_N\right] = \lambda_N u_N. \tag{48}$$

This implies that the sequence $(u_N)_{N\in\mathbb{N}}$, which is uniformly bounded in $H^1_\#(\Omega)$, is in fact also uniformly bounded in $H^{\sigma+2}_\#(\Omega)$. Together with (9), this implies in turn that $(w_{u,u_N})_{N\in\mathbb{N}}$ is bounded in $H^1_\#(\Omega) \cap L^\infty(\Omega)$. We therefore obtain from (47) that

$$|\lambda_N - \lambda| \le C\left(\|u_N - u\|^2_{H^1} + \|u_N - u\|_{H^{-1}}\right). \tag{49}$$

Let us now compute the H^{-1}-norm of the error. Let $w \in H^1_\#(\Omega)$. Proceeding as in Section 2, we obtain

$$\begin{aligned}
&\int_\Omega w(u_N - u)\\
&\quad= -\int_\Omega \left(f(u_N^2)u_N - f(u^2)u_N - 2f'(u^2)u^2(u_N - u)\right)\Pi_N\psi_w\\
&\qquad+(\lambda_N - \lambda)\int_\Omega u_N(\Pi_N\psi_w - \psi_w) + (\lambda_N - \lambda)\int_\Omega (u_N - u)\psi_w\\
&\qquad+\langle (E''(u) - \lambda)(u_N - u), \psi_w - \Pi_N\psi_w\rangle_{X',X}\\
&\qquad+\|u_N - u\|^2_{L^2}\int_\Omega f'(u^2)u^3\psi_w - \frac{1}{2}\|u_N - u\|^2_{L^2}\int_\Omega uw.
\end{aligned}$$

Combining (19), (43), (44)-(46), (47) and the above equality, we obtain that there exists a constant $C \in \mathbb{R}_+$ such that for all $N \in \mathbb{N}$ and all $w \in H^1_\#(\Omega)$,

$$\begin{aligned}\int_\Omega w(u_N - u) &\leq C' \left(\|u_N - u\|^2_{H^1} + N^{-2} \|u_N - u\|_{H^1} \right) \|w\|_{H^1} \\ &\leq \frac{C}{N^{\min(2(\sigma+1),\sigma+3)}} \|w\|_{H^1}.\end{aligned}$$

Therefore

$$\|u_N - u\|_{H^{-1}} = \sup_{w \in H^1_\#(\Omega)\setminus\{0\}} \frac{\displaystyle\int_\Omega w(u_N - u)}{\|w\|_{H^1}} \leq \frac{C}{N^{\min(2(\sigma+1),\sigma+3)}}, \tag{50}$$

for some constant $C \in \mathbb{R}_+$ independent of N. We end up with

$$|\lambda_N - \lambda| \leq \frac{C}{N^{\min(2(\sigma+1),\sigma+3)}}.$$

To proceed further, we need to make additional assumptions on the function f. Indeed, if we had a uniform bound on w_{u,u_N} in $H^r_\#(\Omega)$ for some $r > 1$, the above argument would lead to

$$\begin{aligned}|\lambda_N - \lambda| &\leq C \left(\|u_N - u\|^2_{H^1} + \|u_N - u\|_{H^{-r}} \right) \\ &\leq \frac{C}{N^{\min(2(\sigma+1),\sigma+r+2)}}.\end{aligned}$$

A simple case when this is true is when $F(t) = ct^2$ with $c > 0$; in this particular case indeed, $w_{u,u_N} = 2cu_N^2(u_N + u)$ is uniformly bounded in $H^{\sigma+2}_\#(\Omega)$. We can summarize the results obtained in this section in the following theorem.

Theorem 2 *Assume that $V \in H^\sigma_\#(\Omega)$ for some $\sigma > d/2$ and that the function F satisfies (7)-(9) and is such that the function $t \mapsto F(t^2)$ is in $C^{\sigma+1}(\mathbb{R}_+, \mathbb{R})$ if $\sigma \in \mathbb{N}$, and in $C^{[\sigma]+2}(\mathbb{R}_+, \mathbb{R})$ if $\sigma \notin \mathbb{N}$. Then there exists $C \in \mathbb{R}_+$ such that for all $N \in \mathbb{N}$,*

$$\|u_N - u\|_{H^1} \leq \frac{C}{N^{\sigma+1}} \tag{51}$$

$$\|u_N - u\|_{L^2} \leq \frac{C}{N^{\sigma+2}} \tag{52}$$

$$|\lambda_N - \lambda| \leq \frac{C}{N^{\min(2(\sigma+1),\sigma+3)}}.$$

If in addition

$$w_{u,u_N} := u_N^2 \frac{f(u_N^2) - f(u^2)}{u_N - u}$$

is uniformly bounded in $H^{\sigma+2}_\#(\Omega)$, then

$$|\lambda_N - \lambda| \leq \frac{C}{N^{2(\sigma+1)}}. \tag{53}$$

In order to evaluate the quality of the error bounds obtained in Theorem 2, we have performed numerical tests with $\Omega = (0, 2\pi)$, $V(x) = \sin(|x-\pi|/2)$ and $F(t^2) = t^2/2$. The Fourier coefficients of the potential V are given by

$$\widehat{V}_k = -\frac{1}{\sqrt{2\pi}} \frac{1}{|k|^2 - \frac{1}{4}}, \tag{54}$$

from which we deduce that $V \in H^\sigma_\#(0,2\pi)$ for all $\sigma < 3/2$. It can be see on Figure 1 that $\|u_N - u\|_{H^1}$, $\|u_N - u\|_{L^2}$, $\|u_N - u\|_{H^{-1}}$, and $|\lambda_N - \lambda|$ decay respectively as $N^{-2.67}$, $N^{-3.67}$, $N^{-4.67}$ and N^{-5} (the reference values for u and λ are those obtained for $N = 65$). These results are in very good agreement with the upper bounds (51), (52), (50) and (53), which respectively decay as $N^{-2.5+\epsilon}$, $N^{-3.5+\epsilon}$, $N^{-4.5+\epsilon}$ and $N^{-5+\epsilon}$, for $\epsilon > 0$ arbitrarily small.

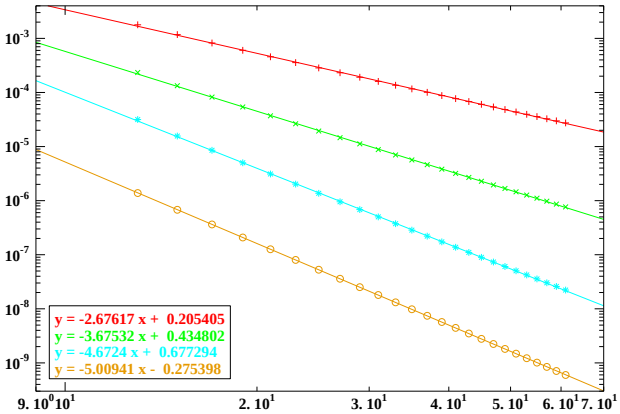

Figure 1: Numerical errors $\|u_N - u\|_{H^1}$ $(+)$, $\|u_N - u\|_{L^2}$ $(\times)$, $\|u_N - u\|_{H^{-1}}$ $(*)$, and $|\lambda_N - \lambda|$ $(\circ)$, as functions of $2N+1$ (the dimension of $\widetilde{X}_N$) in log scales.

4 Finite element discretization

In this section, we consider the problem

$$\inf \left\{ E(v),\ v \in X,\ \int_\Omega v^2 = 1 \right\}, \tag{55}$$

where

$$\begin{aligned} &\Omega \text{ is a rectangular brick of } \mathbb{R}^d, \quad \text{with } d = 1,\ 2 \text{ or } 3, \\ &X = H^1_0(\Omega), \\ &E(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 + \frac{1}{2}\int_\Omega V v^2 + \frac{1}{2}\int_\Omega F(v^2). \end{aligned}$$

We assume that V satisfies (6) and that the function F satisfies (7)-(9). Througout this section, we denote by u the unique positive solution of (55) and by λ the corresponding Lagrange multiplier.

In the non periodic case considered here, a classical variational approximation of (1) is provided by the finite element method. We consider a family of regular triangulations $(\mathcal{T}_h)_h$ of Ω. This means, in the case when $d = 3$ for instance, that for each $h > 0$, $\mathcal{T}_h$ is a collection of tetrahedra such that

- $\overline{\Omega}$ is the union of all the elements of $\mathcal{T}_h$;
- the intersection of two different elements of $\mathcal{T}_h$ is either empty, a vertex, a whole edge, or a whole face of both of them;

- the ratio of the diameter h_K of any element K of $\mathcal{T}_h$ to the diameter of its inscribed sphere is smaller than a constant independent of h.

As usual, h denotes the maximum of the diameters h_K, $K \in \mathcal{T}_h$. The parameter of the discretization then is $\delta = h > 0$. For each K in $\mathcal{T}_h$ and each nonnegative integer k, we denote by $\mathbb{P}_k(K)$ the space of the restrictions to K of the polynomials with d variables and total degree lower or equal to k.

The finite element space $X_{h,k}$ constructed from $\mathcal{T}_h$ and $\mathbb{P}_k(K)$ is the space of all continuous functions on Ω vanishing on $\partial\Omega$ such that their restrictions to any element K of $\mathcal{T}_h$ belong to $\mathbb{P}_k(K)$. Recall that $X_{h,k} \subset H_0^1(\Omega)$ as soon as $k \geq 1$.

We denote by $\pi^0_{h,k}$ and $\pi^1_{h,k}$ the orthogonal projectors on $X_{h,k}$ for the L^2 and H^1 inner products respectively. The following estimates are classical: there exists $C \in \mathbb{R}_+$ such that for all $r \in \mathbb{N}$ such that $1 \leq r \leq k+1$,

$$\forall \phi \in H^r(\Omega) \cap H_0^1(\Omega), \qquad \|\phi - \pi^0_{h,k}\phi\|_{L^2} \leq Ch^r \|\phi\|_{H^r}, \tag{56}$$

$$\forall \phi \in H^r(\Omega) \cap H_0^1(\Omega), \qquad \|\phi - \pi^1_{h,k}\phi\|_{H^1} \leq Ch^{r-1} \|\phi\|_{H^r}. \tag{57}$$

Let $u_{h,k}$ be a solution to the variational problem

$$\inf \left\{ E(v_{h,k}),\ v_{h,k} \in X_{h,k},\ \int_\Omega v_{h,k}^2 = 1 \right\} \tag{58}$$

such that $(u_{h,k}, u)_{L^2} \geq 0$. In this setting, we obtain the following *a priori* error estimates.

Theorem 3 *Assume that V satisfies (6) and that the function F satisfies (7)-(9). Then there exists $h_0 > 0$ and $C \in \mathbb{R}_+$ such that for all $0 < h \leq h_0$,*

$$\|u_{h,1} - u\|_{H^1} \leq C\,h \tag{59}$$

$$\|u_{h,1} - u\|_{L^2} \leq C\,h^2 \tag{60}$$

$$|\lambda_{h,1} - \lambda| \leq C\,h^2. \tag{61}$$

If in addition, $V \in H^1(\Omega)$, then there exists $h_0 > 0$ and $C \in \mathbb{R}_+$ such that for all $0 < h \leq h_0$,

$$\|u_{h,2} - u\|_{H^1} \leq C\,h^2 \tag{62}$$

$$\|u_{h,2} - u\|_{L^2} \leq C\,h^3 \tag{63}$$

$$|\lambda_{h,2} - \lambda| \leq C\,h^4. \tag{64}$$

Proof As Ω is a convex polygon, and as V and F satisfie (6) and (7)-(9) respectively, we have $u \in H^2(\Omega)$. We then use the fact that $\psi_{u_{h,k}-u}$ is solution to

$$\begin{aligned} &-\Delta\psi_{u_{h,k}-u} + (V + f(u^2) + 2f'(u^2)u^2 - \lambda)\psi_{u_{h,k}-u} \\ &\qquad = 2\left(\int_\Omega f'(u^2)u^3\psi_{u_{h,k}-u}\right)u + (u_{h,k} - u) - (u_{h,k} - u, u)_{L^2}u, \end{aligned} \tag{65}$$

to establish that $\psi_{u_{h,k}-u} \in H^2(\Omega) \cap H_0^1(\Omega)$ and that

$$\|\psi_{u_{h,k}-u}\|_{H^2} \leq C\|u_{h,k} - u\|_{L^2} \tag{66}$$

for some constant C independent of h and k. The estimates (59), (60) and (61) then are directly consequences of Theorem 1, (57) and (66).

Under the additional assumption that $V \in H^1(\Omega)$, we obtain by elliptic regularity arguments that $u \in H^3(\Omega)$. The H^1 and L^2 estimates (62) and (63) immediately follows from Theorem 1, (57) and (66). We also have

$$|\lambda_{2,h} - \lambda| \leq Ch^3 \tag{67}$$

for a constant C independent of h. In order to prove (64), we proceed as in Section 3. We start from the equality

$$\lambda_{2,h} - \lambda = \langle (A_u - \lambda)(u_{2,h} - u), (u_{2,h} - u) \rangle_{X',X} + \int_\Omega \widetilde{w}^h (u_{2,h} - u) \tag{68}$$

where

$$\widetilde{w}^h = u_{2,h}^2 \frac{f(u_{2,h}^2) - f(u^2)}{u_{2,h} - u}.$$

We now claim that $u_{h,2}$ converges to u in $L^\infty(\Omega)$ when h goes to zero. To establish this result, we first remark that

$$\|u_{h,2} - u\|_{L^\infty} \leq \|u_{h,2} - \mathcal{I}_{h,2}u\|_{L^\infty} + \|\mathcal{I}_{h,2}u - u\|_{L^\infty},$$

where $\mathcal{I}_{h,2}$ is the interpolation projector on $X_{h,2}$. As $u \in H^3(\Omega) \hookrightarrow C^1(\overline{\Omega})$, we have

$$\lim_{h \to 0^+} \|\mathcal{I}_{h,2}u - u\|_{L^\infty} = 0.$$

On the other hand, using the inverse inequality

$$\exists C \in \mathbb{R}_+ \text{ s.t. } \forall 0 < h \leq h_0, \ \forall v_h \in X_{h,2}, \quad \|v_{h,2}\|_{L^\infty} \leq C\rho(h)\|v_{h,2}\|_{H^1},$$

with $\rho(h) = 1$ if $d = 1$, $\rho(h) = 1 + \ln h$ if $d = 2$ and $\rho(h) = h^{-1/2}$ if $d = 3$ (see [8] for instance), we obtain

$$\begin{aligned} \|u_{h,2} - \mathcal{I}_{h,2}u\|_{L^\infty} &\leq C\rho(h)\|u_{h,2} - \mathcal{I}_{h,2}u\|_{H^1} \\ &\leq C\rho(h)\left(\|u_{h,2} - u\|_{H^1} + \|u - \mathcal{I}_{h,2}u\|_{H^1}\right) \\ &\leq C'\,\rho(h)\,h^2 \underset{h \to 0^+}{\longrightarrow} 0. \end{aligned}$$

Hence the announced result. This implies in particular that $\widetilde{w}^h$ is bounded in $H^1(\Omega)$, uniformly in h. Consequently, there exists $C \in \mathbb{R}_+$ such that for all $0 < h \leq h_0$,

$$|\lambda_{h,2} - \lambda| \leq C\left(\|u_{h,2} - u\|_{H^1}^2 + \|u_{h,2} - u\|_{H^{-1}}\right). \tag{69}$$

To estimate the H^{-1}-norm of $u_{h,2} - u$, we write that for all $w \in H_0^1(\Omega)$,

$$\begin{aligned} &\int_\Omega w(u_{h,2} - u) \\ &\quad = -\int_\Omega \left(f(u_{h,2}^2)u_{h,2} - f(u^2)u_{h,2} - 2f'(u^2)u^2(u_{h,2} - u)\right)\pi_{h,2}^1\psi_w \\ &\qquad +(\lambda_N - \lambda)\int_\Omega u_N(\pi_{h,2}^1\psi_w - \psi_w) + (\lambda_{h,2} - \lambda)\int_\Omega (u_{h,2} - u)\psi_w \\ &\qquad +\langle (E''(u) - \lambda)(u_{h,2} - u), \psi_w - \pi_{h,2}^1\psi_w \rangle_{X',X} \\ &\qquad +\|u_{h,2} - u\|_{L^2}^2 \int_\Omega f'(u^2)u^3\psi_w - \frac{1}{2}\|u_{h,2} - u\|_{L^2}^2 \int_\Omega uw, \end{aligned} \tag{70}$$

where ψ_w is solution to

$$\begin{aligned} &-\Delta\psi_w + (V + f(u^2) + 2f'(u^2)u^2 - \lambda)\psi_w \\ &\qquad = 2\left(\int_\Omega f'(u^2)u^3\psi_w\right)u + w - (w,u)_{L^2}u. \end{aligned} \tag{71}$$

It then follows from (71) that ψ_w is in $H^3(\Omega)$ and that there exists $C \in \mathbb{R}_+$ such that for all $w \in H^1_0(\Omega)$ and all $0 < h \leq h_0$,

$$\|\psi_w\|_{H^3} \leq C\|w\|_{H^1}.$$

We therefore obtain the inequality

$$\|\psi_w - \pi^1_{h,2}\psi_w\|_{H^1} \leq Ch^2\|w\|_{H^1}, \tag{72}$$

where the constant C is independent of h.

Putting together (9), (19), (57), (62), (63), (67) and (72), we get

$$\|u_{h,2} - u\|_{H^{-1}} = \sup_{w \in H^1_0(\Omega)\setminus\{0\}} \frac{\int_\Omega w(u_{h,2} - u)}{\|w\|_{H^1}} \leq C\, h^4.$$

Combining with (62) and (69), we end up with (64). □

Numerical results for the case when $\Omega = (0,\pi)^2$, $V(x_1,x_2) = x_1^2 + x_2^2$ and $F(t^2) = t^2/2$ are reported on Figure 2. The agreement with the error estimates obtained in Theorem 3 is good for the $\mathbb{P}_1$ approximation and excellent for the $\mathbb{P}_2$ approximation.

5 The effect of numerical integration

Let us now address one further consideration that is related to the practical implementation of the method, and more precisely to the numerical integration of the nonlinear term. For simplicity, we focus on the case when $A = 1$.

From a practical viewpoint, the solution $(u_\delta, \lambda_\delta)$ to the nonlinear eigenvalue problem (15) can be computed iteratively, using for instance the optimal damping algorithm [4, 2, 7]. At the p^{th} iteration ($p \geq 1$), the ground state $(u^p_\delta, \lambda^p_\delta) \in X_\delta \times \mathbb{R}$ of some *linear*, finite dimensional, eigenvalue problem of the form

$$\int_\Omega \overline{\nabla u^p_\delta} \cdot \nabla v_\delta + \int_\Omega \left(V + f(\widetilde{\rho}^{p-1}_\delta)\right) \overline{u^p_\delta}\, v_\delta = \lambda^p_\delta \int_\Omega \overline{u^p_\delta} v_\delta, \quad \forall v_\delta \in X_\delta, \tag{73}$$

has to be computed. In the optimal damping algorithm, the density $\widetilde{\rho}^{p-1}_\delta$ is a convex linear combination of the densites $\rho^q_\delta = |u^q_\delta|^2$, for $0 \leq q \leq p-1$. Solving (73) amounts to finding the lowest eigenelement of the matrix H^p with entries

$$H^p_{kl} := \int_\Omega \overline{\nabla \phi_k} \cdot \nabla \phi_l + \int_\Omega V\, \overline{\phi_k}\, \phi_l + \int_\Omega f(\widetilde{\rho}^{p-1}_\delta)\, \overline{\phi_k}\, \phi_l, \tag{74}$$

where $(\phi_k)_{1 \leq k \leq \dim(X_\delta)}$ stands for the canonical basis of X_δ.

In order to evaluate the last two terms of the right-hand side of (74), numerical integration has to be resorted to. In the finite element approximation of (55), it is generally made use of a numerical quadrature formula over each triangle (2D) or tetrahedron (3D) based on Gauss points. In the Fourier approximation of the periodic problem (39), the terms

$$\int_\Omega V\, \overline{e_k}\, e_l \quad \text{and} \quad \int_\Omega f(\widetilde{\rho}^{p-1}_\delta)\, \overline{e_k}\, e_l,$$

which are in fact, up to a multiplicative constant, the $(k-l)^{\text{th}}$ Fourier coefficients of V and $f(\widetilde{\rho}^{p-1}_\delta)$ respectively, are evaluated by Fast Fourier Transform (FFT), using an integration grid which may be different from the natural discretization grid

$$\left\{ \left(\frac{2\pi}{2N+1} j_1, \cdots, \frac{2\pi}{2N+1} j_d, \right),\ 0 \leq j_1, \cdots, j_d \leq 2N \right\}$$

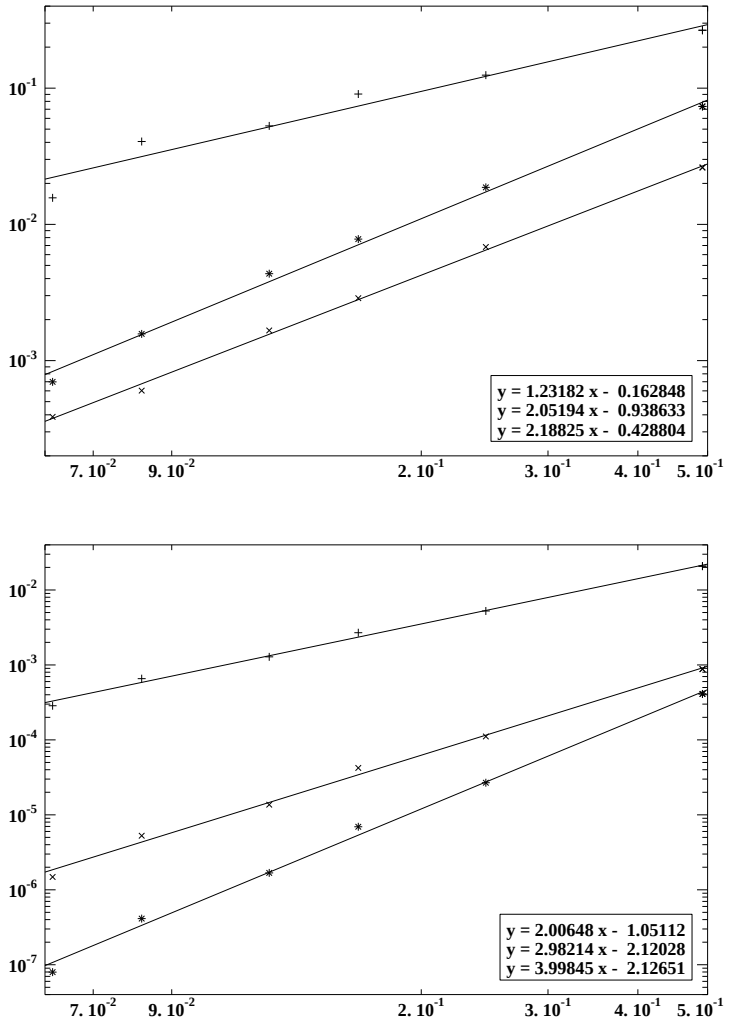

Figure 2: Errors $\|u_{h,k} - u\|_{H^1}$ $(+)$, $\|u_{h,k} - u\|_{L^2}$ $(\times)$ and $|\lambda_{h,k} - \lambda|$ $(*)$ for the $\mathbb{P}_1$ $(k = 1$, top) and $\mathbb{P}_2$ $(k = 2$, bottom) approximations as a function of h in log scales.

associated with $\widetilde{X}_N$. This raises the question of the influence of the numerical integration on the convergence results obtained in Theorems 1, 2 and 3.

Remark 2 *In the case of the periodic problem considered in Section 3 and when $F(t) = ct^2$ for some $c > 0$, the last term of the right-hand side of (74) can be computed exactly (up to round-off errors) by means of a Fast Fourier Transform (FFT) on an integration grid twice as fine as the discretization grid. This is due to the fact that the function $\widetilde{\rho}_\delta^{p-1}\,\overline{e_k}\,e_l$ belongs to the space Span$\{e_n \,|\, |n|_* \leq 4N\}$. An analogous property is used in the evaluation of the Coulomb term in the numerical simulation of the Kohn-Sham equations for periodic systems.*

In the sequel, we focus on the simple case when $d = 1$, $\Omega = [0, 2\pi)$, $X = H^1_\#(0, 2\pi)$, and

$$E(v) = \frac{1}{2}\int_0^{2\pi} |v'|^2 + \frac{1}{2}\int_0^{2\pi} Vv^2 + \frac{1}{4}\int_0^{2\pi} |v|^4$$

with $V \in H^\sigma_\#(0,2\pi)$ for some $\sigma > 1/2$. More difficult cases will be addressed elsewhere [5].

In view of Remark 2, we consider an integration grid

$$\frac{2\pi}{N_g}\mathbb{Z} \cap [0,2\pi) = \left\{0, \frac{2\pi}{N_g}, \frac{4\pi}{N_g}, \cdots, \frac{2\pi(N_g-1)}{N_g}\right\},$$

with $N_g \geq 4N+1$ for which we have

$$\forall v_N \in \widetilde{X}_N, \quad \int_0^{2\pi} |v_N|^4 = \frac{2\pi}{N_g} \sum_{r \in \frac{2\pi}{N_g}\mathbb{Z}\cap[0,2\pi)} |v_N(r)|^4,$$

and for all $\rho \in \widetilde{X}_{2N}$,

$$\forall |k|,|l| \leq N, \quad \int_0^{2\pi} \rho\, \overline{e_k}\, e_l = \frac{1}{N_g} \sum_{r \in \frac{2\pi}{N_g}\mathbb{Z}\cap[0,2\pi)} \rho(r) e^{-i(k-l)r} = \widehat{\rho^{\rm FFT}_{k-l}}, \tag{75}$$

where $\widehat{\rho^{\rm FFT}_{k-l}}$ is the $(k-l)^{\rm th}$ coefficient of the discrete Fourier transform of ρ. Recall that if $\phi = \sum_{g\in\mathbb{Z}} \widehat{\phi}_g\, e_g \in C^0_\#(0,2\pi)$, the discrete Fourier transform of ϕ is the $N_g\mathbb{Z}$-periodic sequence $(\widehat{\phi^{\rm FFT}_g})_{g\in\mathbb{Z}}$ defined by

$$\forall g \in \mathbb{Z}, \quad \widehat{\phi^{\rm FFT}_g} = \frac{1}{N_g} \sum_{r \in \frac{2\pi}{N_g}\mathbb{Z}\cap[0,2\pi)} \phi(r) e^{-igr}.$$

We now introduce the subspaces W_M for $M \in \mathbb{N}^*$ such that $W_M = \widetilde{X}_{(M-1)/2}$ if M is odd and $W_M = \widetilde{X}_{M/2-1} \oplus \mathbb{C}(e_{M/2} + e_{-M/2})$ is M is even (note that $\dim(W_M) = M$ for all $M \in \mathbb{N}^*$). It is then possible to define an interpolation projector $\mathcal{I}_{N_g}$ from $C^0_\#(0,2\pi)$ onto W_{N_g} by

$$\forall x \in \frac{2\pi}{N_g}\mathbb{Z} \cap [0,2\pi), \quad [\mathcal{I}_{N_g}(\phi)](x) = \phi(x).$$

The expansion of $\mathcal{I}_{N_g}(\phi)$ in the canonical basis of W_{N_g} is given by

$$\mathcal{I}_{N_g}(\phi) = \left| \begin{array}{ll} (2\pi)^{1/2} \displaystyle\sum_{|g|\leq (N_g-1)/2} \widehat{\phi^{\rm FFT}_g}\, e_g & (N_g \text{ odd}), \\ (2\pi)^{1/2} \displaystyle\sum_{|g|\leq N_g/2-1} \widehat{\phi^{\rm FFT}_g}\, e_g + (2\pi)^{1/2} \widehat{\phi^{\rm FFT}_{N_g/2}} \left(\frac{e_{N_g/2} + e_{-N_g/2}}{2}\right) & (N_g \text{ even}). \end{array} \right.$$

Under the condition that $N_g \geq 4N+1$, the following property holds: for all $\phi \in C^0_\#(0,2\pi)$,

$$\forall |k|,|l| \leq N, \quad \int_0^{2\pi} \mathcal{I}_{N_g}(\phi)\, \overline{e_k}\, e_l = \widehat{\phi^{\rm FFT}_{k-l}}.$$

It is therefore possible, in the particular case considered here, to efficiently evaluate the entries of the matrix H^p using the formula

$$\begin{aligned} H^p_{kl} &:= \int_0^{2\pi} \overline{e'_k} \cdot e'_l + \int_0^{2\pi} V\, \overline{e_k}\, e_l + \int_0^{2\pi} \widetilde{\rho}_N^{p-1} \overline{e_k}\, e_l \\ &\simeq |k|^2 \delta_{kl} + \widehat{V^{\rm FFT}_{k-l}} + \widehat{[\widetilde{\rho}_N^{p-1}]^{\rm FFT}_{k-l}}, \end{aligned} \tag{76}$$

and resorting to Fast Fourier Transform (FFT) algorithms to compute the discrete Fourier transforms. Note that only the second term is computed approximatively. The third term is computed exactly since, at each iteration, $\widetilde{\rho}_N^{p-1}$ belongs to $\widetilde{X}_{2N}$ (see Eq. (75)). Of course, this situation is specific to the nonlinearity $F(t) = t^2/2$ considered here.

Using the approximation formula (76) amounts to replace the original problem

$$\inf\left\{E(v_N),\ v_N \in \widetilde{X}_N,\ \int_0^{2\pi} |v_N|^2 = 1\right\}, \tag{77}$$

with the approximate problem

$$\inf\left\{E_{N_b}(v_N),\ v_N \in \widetilde{X}_N,\ \int_0^{2\pi} |v_N|^2 = 1\right\}, \tag{78}$$

where

$$E_{N_b}(v_N) = \frac{1}{2}\int_0^{2\pi} |v_N'|^2 + \frac{1}{2}\int_0^{2\pi} \mathcal{I}_{N_g}(V) v_N^2 + \frac{1}{4}\int_0^{2\pi} |v_N|^4.$$

Let us denote by u_N a solution of (77) such that $(u_N, u)_{L^2} \geq 0$ and by u_{N,N_g} a solution to (78) such that $(u_{N,N_g}, u)_{L^2} \geq 0$. It is easy to check that u_{N,N_g} is bounded in $H^1_\#(0,2\pi)$ uniformly in N and N_g.

Besides, we know from Theorem 2 that $(u_N)_{N\in\mathbb{N}}$ converges to u in $H^1_\#(0,2\pi)$, hence in $L^\infty_\#(2,\pi)$, when N goes to infinity. This implies that the sequence $(A_u - A_{u_N})_{N\in\mathbb{N}}$ converges to 0 in operator norm. Consequently, for all N large enough and all N_g such that $N_g \geq 4N+1$,

$$\begin{aligned}
\frac{\gamma}{4}\|u_{N,N_g} - u_N\|_{H^1}^2 &\leq E(u_{N,N_g}) - E(u_N) \\
&\leq E_{N_g}(u_{N,N_g}) - E_{N_g}(u_N) \\
&\quad + \int_0^{2\pi} (V - \mathcal{I}_{N_g}(V))\left(|u_{N,N_g}|^2 - |u_N|^2\right) \\
&\leq \int_0^{2\pi} (V - \mathcal{I}_{N_g}(V))\left(|u_{N,N_g}|^2 - |u_N|^2\right) \\
&\leq C\|\Pi_{2N}(V - \mathcal{I}_{N_g}(V))\|_{L^2}\|u_{N,N_g} - u_N\|_{H^1},
\end{aligned}$$

where we have used the fact that $\left(|u_{N,N_g}|^2 - |u_N|^2\right) \in \widetilde{X}_{2N}$. Therefore,

$$\|u_{N,N_g} - u_N\|_{H^1} \leq C\|\Pi_{2N}(V - \mathcal{I}_{N_g}(V))\|_{L^2}, \tag{79}$$

for a constant C independent of N and N_g. Likewise,

$$\begin{aligned}
\lambda_{N,N_g} - \lambda_N &= \langle (A_{u_N} - \lambda_N)(u_{N,N_g} - u_N), (u_{N,N_g} - u_N)\rangle_{X',X} \\
&\quad + \int_0^{2\pi} (V - \mathcal{I}_N(V))|u_{N,N_g}|^2 \\
&\quad + \int_0^{2\pi} |u_{N,N_g}|^2 (u_{N,N_g} + u_N)(u_{N,N_g} - u_N),
\end{aligned}$$

from which we deduce, using (79),

$$|\lambda_{N,N_g} - \lambda_N| \leq C\|\Pi_{2N}(V - \mathcal{I}_{N_g}(V))\|_{L^2}.$$

An error analysis of the interpolation operator $\mathcal{I}_{N_g}$ is given in [6]: for all non-negative real numbers $0 \leq r \leq s$ with $s > 1/2$ (for $d = 1$),

$$\|\varphi - \mathcal{I}_{N_g}(\varphi)\|_{H^r} \leq \frac{C}{N_g^{s-r}}\|\varphi\|_{H^s}, \quad \forall \varphi \in H^s_\#(0,2\pi). \tag{80}$$

Thus,

$$\|\Pi_{2N}(V - \mathcal{I}_{N_g}(V))\|_{L^2} \leq \|V - \mathcal{I}_{N_g}(V)\|_{H^\sigma} \leq \frac{C}{N_g^\sigma}, \tag{81}$$

and the above inequality provides the following estimates:

$$\begin{aligned} \|u_{N,N_g} - u\|_{H^1} &\leq C\left(N^{-\sigma-1} + N_g^{-\sigma}\right) && (82)\\ \|u_{N,N_g} - u\|_{L^2} &\leq C\left(N^{-\sigma-2} + N_g^{-\sigma}\right) && (83)\\ |\lambda_{N,N_g} - \lambda| &\leq C\left(N^{-2\sigma-2} + N_g^{-\sigma}\right), && (84) \end{aligned}$$

for a constant C independent of N and N_g. The first component of the error bound (82) corresponds to the error $\|u_N - u\|_{H^1}$ while the second component corresponds to the numerical integration error $\|u_{N,N_g} - u_N\|_{H^1}$ (the same remark applies to the error bounds (83) and (84)).

It is classical that for the norm $\|\varphi - \mathcal{I}_{N_g}\varphi\|_{H^r}$ for $r < 0$ is in general of the same order of magnitude as $\|\varphi - \mathcal{I}_{N_g}\varphi\|_{L^2}$. As the existence of better estimates in negative norms is a corner stone in the derivation of the improvement of the error estimate (46) for the eigenvalues (doubling of the convergence rate), we expect that the eigenvalue approximation will be dramatically polluted by the use of the numerical integration formula.

This can be checked numerically. Considering again the one-dimensional example used in Section 3 ($\Omega = (0, 2\pi)$, $V(x) = \sin(|x - \pi|/2)$, $F(t) = t^2/2$), we have computed for $4 \leq N \leq 30$ and $N_g = 2^p$ with $7 \leq p \leq 15$, the errors $\|u_{N,N_g} - u\|_{H^1}$, $\|u_{N,N_g} - u\|_{L^2}$, $\|u_{N,N_g} - u\|_{H^{-1}}$, and $|\lambda_{N,N_g} - \lambda|$. On Figure 3, these quantities are plotted as functions of $2N + 1$ (the dimension of $\widetilde{X}_N$), for various values of N_g.

The non-monotonicity of the curve $N \mapsto |\lambda_{N,N_g} - \lambda|$ originates from the fact that $\lambda_{N,N_g} - \lambda$ can be positive or negative depending on the values of N and N_g.

The numerical errors $\|u_{N,N_g} - u\|_{H^1}$, $\|u_{N,N_g} - u\|_{L^2}$, $\|u_{N,N_g} - u\|_{H^{-1}}$, and $|\lambda_{N,N_g} - \lambda|$, for $N = 30$, as functions of N_g (in log scales) are plotted on Figure 4. When N_g goes to infinity, the sequences $\log_{10}\|u_{N,N_g} - u\|_{H^1}$, $\log_{10}\|u_{N,N_g} - u\|_{L^2}$, $\log_{10}\|u_{N,N_g} - u\|_{H^{-1}}$, and $\log_{10}|\lambda_{N,N_g} - \lambda|$ converge to $\log_{10}\|u_N - u\|_{H^1}$, $\log_{10}\|u_N - u\|_{L^2}$, $\log_{10}\|u_N - u\|_{H^{-1}}$, and $\log_{10}|\lambda_N - \lambda|$ respectively. For smaller values of N_g, the numerical integration error dominates and these functions all decay linearly with $\log_{10} N_g$ with a slope very close to -2. For fixed N, the upper bounds (82)-(84) also decay linearly with $\log_{10} N_g$, but with a slope equal to -1.5. To obtain sharper upper bounds for the numerical integration error, we need to replace (81) with a sharper estimate of $\|\Pi_{2N}(V - \mathcal{I}_{N_g}(V))\|_{L^2}$, which is possible for the particular example under consideration here. Indeed, remarking that under the condition $N_g \geq 4N + 1$,

$$\|\Pi_{2N}(V - \mathcal{I}_{N_g}(V))\|_{L^2} = \left(\sum_{|g|\leq 2N}\left|\sum_{k\in\mathbb{Z}^*}\widehat{V}_{g+kN_g}\right|^2\right)^{1/2},$$

we can, using (54), show that

$$\|\Pi_{2N}(V - \mathcal{I}_{N_g}(V))\|_{L^2} \leq \frac{C\,N^{1/2}}{N_g^2},$$

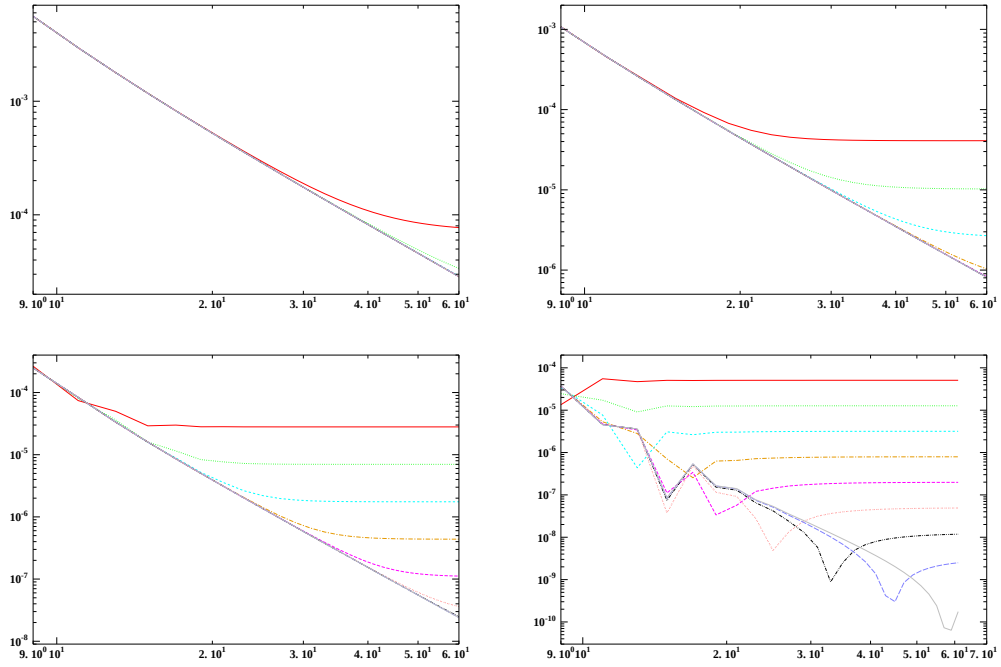

Figure 3: Numerical errors $\|u_{N,N_g} - u\|_{H^1}$ (top left), $\|u_{N,N_g} - u\|_{L^2}$ (top right), $\|u_{N,N_g} - u\|_{H^{-1}}$ (bottom left), and $|\lambda_{N,N_g} - \lambda|$ (bottom right), as functions of $2N+1$ (the dimension of $\widetilde{X}_N$), for $N_g = 128$ (red), $N_g = 256$ (green), $N_g = 512$ (cyan), $N_g = 1024$ (gold), $N_g = 2048$ (magenta), $N_g = 4096$ (pink), $N_g = 8192$ (black), $N_g = 16384$ (blue), $N_g = 32768$ (light blue).

for a constant C independent of N and N_g. We deduce that for this specific example

$$\begin{aligned} \|u_{N,N_g} - u\|_{H^1} &\leq C\left(N^{-5/2} + N^{1/2}N_g^{-2}\right) \\ \|u_{N,N_g} - u\|_{L^2} &\leq C\left(N^{-7/2} + N^{1/2}N_g^{-2}\right) \\ |\lambda_{N,N_g} - \lambda| &\leq C\left(N^{-9/2} + N^{1/2}N_g^{-2}\right). \end{aligned}$$

Acknowledgements

This work was done while E.C. was visiting the Division of Applied Mathematics of Brown University, whose support is gratefully acknowledged. The authors also thank Didier Smets for fruitful discussions.

6 Appendix: properties of the ground state

The mathematical properties of the minimization problems (1) and (10) which are useful for the numerical analysis reported in this article are gathered in the following lemma.

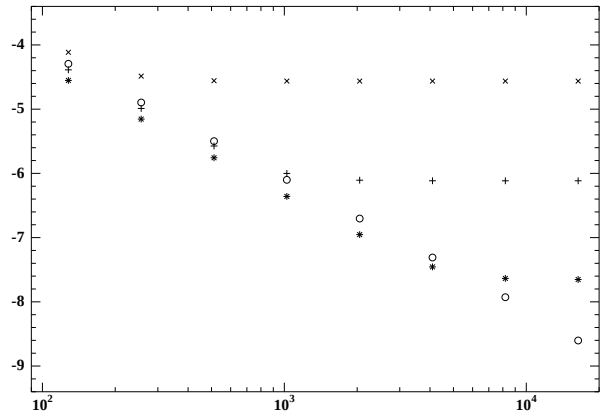

Figure 4: Numerical errors $\|u_{N,N_g} - u\|_{H^1}$ ($\times$), $\|u_{N,N_g} - u\|_{L^2}$ ($+$), $\|u_{N,N_g} - u\|_{H^{-1}}$ ($*$), and $|\lambda_{N,N_g} - \lambda|$ ($\circ$), for $N = 30$, as functions of N_g (in log scales).

Lemma 2 *Under assumptions (4)-(9), (10) has a unique minimizer ρ_0 and (1) has exactly two minimizers $u = \sqrt{\rho_0}$ and $-u$. The function u is solution to the nonlinear eigenvalue problem (12) for some $\lambda \in \mathbb{R}$. Besides, $u \in C^{0,\alpha}(\overline{\Omega})$ for some $0 < \alpha < 1$, $u > 0$ in Ω, and λ is the lowest eigenvalue of A_u and is non-degenerate.*

Proof As A is uniformly bounded and coercive on Ω and $V \in L^q(\Omega)$ for some $q > 2$, $v \mapsto a(v,v)$ is a quadratic form on X, bounded below on the set $\{v \in X \mid \|v\|_{L^2} = 1\}$. Replacing $a(v,v)$ with $a(v,v) - C\|v\|_{L^2}^2$ and $F(t)$ with $F(t) - F(0) - tF'(0)$ does not change the minimizers of (1) and (10). We can therefore assume, *without loss of generality*, that

$$\forall v \in X,\ a(v,v) \geq \|v\|_{L^2}^2 \quad \text{and} \quad F(0) = F'(0) = 0. \tag{85}$$

It then follows from (9), making t_1 equal to zero, that $0 \leq F(v^2) \leq C(v^2 + |v|^5)$. As $X \hookrightarrow L^2(\Omega) \cap L^6(\Omega)$, $E(v)$ is finite for all $v \in X$, $I > -\infty$ and the minimizing sequences of (1) are bounded in X. Let $(v_n)_{n\in\mathbb{N}}$ be a minimizing sequence of (1). Using the fact that X is compactly embedded in $L^2(\Omega)$, we can extract from $(v_n)_{n\in\mathbb{N}}$ a subsequence $(v_{n_k})_{k\in\mathbb{N}}$ which converges weakly in X, strongly in $L^2(\Omega)$ and almost everywhere in Ω to some $u \in X$. As $\|v_{n_k}\|_{L^2} = 1$ and $E(v_{n_k}) \downarrow I$, we obtain $\|u\|_{L^2} = 1$ and $E(u) \leq I$ (E is convex and strongly continuous, hence weakly l.s.c., on X). Hence u is a minimizer of (1). As $|u| \in X$, $\||u|\|_{L^2} = 1$ and $E(|u|) = E(u)$, we can assume without loss of generality that $u \geq 0$. Assumptions (4)-(9) imply that E is C^1 on X and that $E'(u) = A_u u$. It follows that u is solution to (11) for some $\lambda \in \mathbb{R}$. By elliptic regularity arguments [10], we get $u \in C^{0,\alpha}(\overline{\Omega})$ for some $0 < \alpha < 1$. We also have $u > 0$ in Ω; this is a consequence of the Harnack inequality [11]. Making the change of variable $\rho = v^2$, it is easily seen that if v is a minimizer of (1), then v^2 is a minimizer of (10), and that, conversely, if ρ is a minimizer of (10), then $\sqrt{\rho}$ and $-\sqrt{\rho}$ are minimizers of (1). Besides, the functional $\mathcal{E}$ is strictly convex on the convex set $\left\{\rho \geq 0 \mid \sqrt{\rho} \in X,\ \int_\Omega \rho = 1\right\}$. Therefore $\rho_0 = u^2$ is the unique minimizer of (10) and u and $-u$ are the only minimizers of (1).

It is easy to see that A_u is bounded below and has a compact resolvent. It therefore possesses a lowest eigenvalue λ_0, which, according to the min-max principle, satisfies

$$\lambda_0 = \inf\left\{\int_\Omega (A\nabla v)\cdot\nabla v + \int_\Omega (V + f(u^2))v^2,\ v \in X,\ \int_\Omega v^2 = 1\right\}. \tag{86}$$

Let v_0 be a normalized eigenvector of A_u associated with λ_0. Clearly, v_0 is a minimizer of (86) and so is $|v_0|$. Therefore, $|v_0|$ is solution to the Euler equation $A_u|v_0| = \lambda_0|v_0|$. Using again elliptic regularity arguments and the Harnack inequality, we obtain that $|v_0| \in C^{0,\alpha}(\overline{\Omega})$ for some $0 < \alpha < 1$ and that $|v_0| > 0$ on Ω. This implies that either $v_0 = |v_0| > 0$ in Ω or $v_0 = -|v_0| < 0$ in Ω. In particular $(u, v_0)_{L^2} \neq 0$. Consequently, $\lambda = \lambda_0$ and λ is a simple eigenvalue of A_u. □

It is interesting to note that λ is also the ground state eigenvalue of the *nonlinear* eigenvalue problem

$$\left\{ \begin{array}{l} \text{search } (\mu, v) \in \mathbb{R} \times X \text{ such that} \\ A_v v = \mu v \\ \|v\|_{L^2} = 1, \end{array} \right. \tag{87}$$

in the following sense: if (μ, v) is solution to (87) then either $\mu > \lambda$ or $\mu = \lambda$ and $v = \pm u$.

To see this, let us consider a solution $(\mu, v) \in \mathbb{R} \times X$ to (87) and denote by $\widetilde{w} = |v| - u$. As for u, we infer from elliptic regularity arguments [10] that $v \in C^{0,\alpha}(\overline{\Omega})$. We have $\|v\|_{L^2} = \|u\|_{L^2} = 1$. Therefore, if $w \leq 0$ in Ω, then $|v| = u$, which yields $v = \pm u$ and $\mu = \lambda$. Otherwise, there exists $x_0 \in \Omega$ such that $\widetilde{w}(x_0) > 0$, and, up to replacing v with $-v$, we can consider that the function $w = v - u$ is such that $w(x_0) > 0$. The function w is in $X \cap C^{0,\alpha}(\overline{\Omega})$ and satisfies

$$(A_u - \lambda)w + \frac{f(v^2) - f(u^2)}{v^2 - u^2} v(u + v)w = (\mu - \lambda)v. \tag{88}$$

Let $\omega = \{x \in \Omega \mid w(x) > 0\} = \{x \in \Omega \mid v(x) > u(x)\}$ and $w_+ = \max(w, 0)$. As $w_+ \in X$, we deduce from (88) that

$$\langle (A_u - \lambda)w_+, w_+ \rangle_{X',X} + \int_\omega \frac{f(v^2) - f(u^2)}{v^2 - u^2} v(u + v)w = (\mu - \lambda) \int_\omega vw.$$

The left hand side of the above equality is positive and $\int_\omega vw > 0$. Therefore, $\mu > \lambda$.

References

[1] I. Babuška and J. Osborn, *Eigenvalue problems*, in: Handbook of numerical analysis. Volume II, (North-Holland, 1991) 641-787.

[2] E. Cancès, *SCF algorithms for Kohn-Sham models with fractional occupation numbers*, J. Chem. Phys. 114 (2001) 10616-10623.

[3] E. Cancès, M. Defranceschi, W. Kutzelnigg, C. Le Bris and Y. Maday, *Computational quantum chemistry: a primer*, in: Handbook of numerical analysis. Volume X: special volume: computational chemistry, Ph. Ciarlet and C. Le Bris eds (North-Holland, 2003) 3-270.

[4] E. Cancès and C. Le Bris, *Can we outperform the DIIS approach for electronic structure calculations?*, Int. J. Quantum Chem. 79 (2000) 82-90.

[5] E. Cancès, R. Chakir and Y. Maday, *Numerical analysis of the planewave discretization of Kohn-Sham and related models*, in preparation.

[6] C. Canuto, M.Y. Hussaini, A. Quarteroni and T.A. Zang, *Spectral methods*, Springer, 2007.

[7] C. Dion and E. Cancès, *Spectral method for the time-dependent Gross-Pitaevskii equation with harmonic traps*, Phys. Rev. E 67 (2003) 046706.

[8] A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, Springer, 2004.

[9] L. P. Pitaevskii and S. Stringari, *Bose-Einstein condensation*, Clarendon Press, 2003.

[10] D. Gilbarg and N.S. Trudinger, *Elliptic partial differential equations of second order*, 3rd edition, Springer 1998.

[11] G. Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinues*, Ann. Inst. Fourier, tome 15 (1965) 189-257.